

1 **DIRECT TESTIMONY**

2 **OF**

3 **R. DOW BAILEY**

4 **ON BEHALF OF**

5 **SOUTH CAROLINA ELECTRIC & GAS COMPANY**

6 **DOCKET NO. 2003-2-E**

7 **Q. PLEASE STATE YOUR NAME AND BUSINESS ADDRESS.**

8 A. R. Dow Bailey, 1426 Main Street, Columbia, South Carolina.

9 **Q. BY WHOM ARE YOU EMPLOYED AND IN WHAT CAPACITY?**

10 A. I am Forecast Coordinator in the Resource Planning Department of SCANA
11 Corporation.

12 **Q. PLEASE DESCRIBE YOUR EDUCATIONAL BACKGROUND AND BUSINESS**
13 **EXPERIENCE.**

14 A. I am a graduate of Emory University in Atlanta, Georgia where I majored in history. I
15 also received an MBA from the University of Georgia, with an emphasis on finance and
16 economics. I have completed all the coursework requirements for a Ph.D. in economics at
17 the University of South Carolina. In addition to these academic studies I have attended
18 numerous seminars on forecasting and statistics, sponsored by such organizations as
19 NARUC, DOE, the Electric Power Research Institute (EPRI), the National Association of
20 Business Economists (NABE), and the American Gas Association (AGA). Prior to my
21 employment with SCE&G I was employed as an Economic Analyst with Gulf Oil
22 Corporation; an Economist with Wilbur Smith & Associates; a Research Analyst with the
23 South Carolina Public Service Commission; an Economist with CH2M Hill, a consulting

1 engineering firm; and a Financial Analyst with Northeast Utilities. In June 1983 I began
2 work at SCE&G as an Associate Analyst in the Forecasting Department, where I have been
3 employed for the past nineteen years.

4 **Q. WILL YOU BRIEFLY SUMMARIZE YOUR DUTIES WITH SOUTH**
5 **CAROLINA ELECTRIC & GAS COMPANY?**

6 **A.** I am currently responsible for preparing SCE&G's electric and gas forecasts of sales,
7 customers, revenues, and peak demand , as well as other forecasting duties within SCANA.

8 **Q. WHAT IS THE PURPOSE OF YOUR TESTIMONY?**

9 **A.** The purpose of my testimony is to discuss the construction of the short-run electric sales
10 forecast. Short-run is defined as two years, in this case 2003 and 2004.

11 **Q. PLEASE DESCRIBE HOW THE SALES FORECAST WAS DETERMINED.**

12 **A.** The first step in the process was to develop a data set containing monthly information on
13 SCE&G customers, sales, and average use at a class and/or rate level. A number of large
14 industrial customers were separated into individual accounts, as were municipal and
15 cooperative customers. The size of these customers justified developing projections for
16 them individually. Within the residential sector, customers were grouped into three major
17 categories: single-family, multi-family, and mobile home, and further disaggregated into
18 space and non-space heat groups . This is important because there are distinct differences in
19 usage within these housing and heating types that can only be modeled correctly by
20 separating them from each other.

21 At the same time daily weather data from Charleston and Columbia was also
22 gathered for the same historic period. The bulk of SCE&G's residential and commercial
23 customers are located in these two cities, so using an average of the two makes estimation of

1 weather-related sales more reliable. The daily sales were then weighted by scheduled billing
2 cycles to create cycle-weighted weather variables. Weather data calculated on a calendar
3 basis does not correlate well with cycle-billed sales, because the latter are spread over two
4 months, rather than one, so this problem is solved by the weighting process. Weather was
5 represented by heating degree days (HDD) or cooling degree days (CDD). HDD and CDD
6 are defined as the absolute value of the difference between the average daily temperature for
7 a given weather station and some base temperature. HDD are zero if the average daily
8 temperature is above the base temperature, with the converse being true for CDD. Positive
9 HDD indicate the need for space heating and positive CDD likewise the need for space
10 cooling. At SCE&G, the CDD and HDD bases were set at 60° and 75°, respectively.
11 Depending on the season, summer or winter, either CDD or HDD were used to represent
12 weather in the forecast. The months of November through April were considered winter
13 months, with May through October defined as summer. After assembling the data, the actual
14 models were developed. Exhibit__(RDB-1) shows the class/rate breakdown used in the
15 forecast. There are 31 class/rate groupings. However, the use of seasonal and individual
16 customer models means that more equations were developed than the number of classes and
17 rates shown in the exhibit. Overall, nearly 100 models were utilized to develop the short-run
18 forecast.

19 **Q. WHAT TYPES OF MODELS WERE USED TO DEVELOP THE FORECAST?**

20 **A.** Two different types of statistically based models were used to create the forecast, time-
21 series and regressions. Time-series models, from the class of models sometimes known as
22 “Box-Jenkins or “ARIMA” models, were developed to project the bulk of the customer
23 projections. Some customer groups, e.g., municipals and cooperatives, had their customer

1 levels set at a fixed amount because the addition of this type of account is generally known
2 well in advance. On the other hand, classes with large numbers of accounts, such as the
3 residential group, increase according to broad economic and demographic trends within
4 SCE&G's service area. In the short-run, these movements are captured quite accurately by
5 time-series models. In a very few cases time-series models were also used to forecast sales
6 or average use for customers whose usage patterns were not very sensitive to weather.

7 Regression models are a means of statistically relating one variable, such as average
8 use, to other "explanatory" variables, such as weather, price, industrial output, etc. To create
9 a forecast via regression models, first an estimated equation is created using historical data.
10 A number of different model specifications are tried before the final equation is selected.
11 Choice of the final equation is based upon the equation's statistical characteristics, such as
12 its consistency with actual data and the statistical significance of model parameters, its
13 ability to capture turning points in the data, and the reasonableness of equation results. The
14 chosen model is then combined with projected values of the explanatory values to derive the
15 forecast of the dependent variable. If the dependent variable was average use for a given
16 category, as was generally the case, this value was then combined with projected customers
17 to derive sales. Exhibit__(RDB-2) lists the methodologies used to forecast the major
18 rate/class groupings.

19 **Q. CAN YOU GIVE AN EXAMPLE OF AN EQUATION USED TO DEVELOP**
20 **THE FORECAST?**

21 **A.** Yes. Exhibit__(RDB-3) shows a graph relating summer residential average use for the
22 single-family space-heating group to CDD, while Exhibit__(RDB-4) contains output
23 describing regression model results for this same category. The graph in Exhibit__(RDB-3)

1 was included because it illustrates the fact that average use seems more responsive to CDD
2 at lower values than at higher values, a somewhat counter-intuitive result. This phenomenon
3 occurs primarily because lower CDD occur in May and October, the beginning and end of
4 the summer season. In May this result can be explained by startup of air-conditioning units
5 by customers, which when aggregated over a group gives rise to a higher than normal
6 response. In October the reverse seems to be true, and at any rate this factor needs to be
7 accounted for in the forecast. To do so CDD were separated into two groups, CDDLO and
8 CDDHI, so separate coefficients could be estimated. These values, shown under the column
9 heading "Parameter Estimates", were 5.40 and 3.05, respectively. The interpretation of the
10 5.40 coefficient is that average use will increase by 5.4KWH for each CDD, as long as the
11 total CDD are less than 100. When CDD are greater than 100, then the response drops to
12 3.05 KWH for each CDD.

13 **Q. WHAT DO THE OTHER VARIABLES IN THE EQUATION INDICATE?**

14 **A.** There are five variables in the equation which begin with "Y" followed by a year and
15 month, e.g., Y8909. These are sometimes called dummy, indicator or binary variables, and
16 are used to account for the impact of unusual events or other factors not explained by
17 standard explanatory variables. For example, the value of -240.04 was estimated by the
18 regression for Y8909. This represents an estimate of reduced average use during the month
19 of September 1989, when Hurricane Hugo occurred. Accounting for this unusual event
20 allows the other variables to be more accurately estimated in the regression process.

21 SEP is another dummy variable which provides an estimate of higher average use in the
22 month of September after the effects of the other explanatory variables are taken into
23 consideration. September is a transition month, but due to cycle billing a large amount of

1 calculated to be 1585.55. Thus the unexplained variability, or average error of the equation,
2 is about 2.4%. A number of other statistics and graphs not shown in the exhibit were also
3 used in the decision to use this particular specification of residential average use.

4 **Q. WHAT IS THE NEXT STEP IN THE FORECAST PROCESS?**

5 **A.** After all the statistical work is completed and the models have been used to generate a
6 preliminary forecast, the projections are then compared to historic actuals for
7 reasonableness. If certain events affecting sales or customer growth in the future are known,
8 but not contained in the model structure, adjustments are also made for these factors. A
9 prime example of this is information gathered by the Industrial Marketing and Economic
10 Development departments on SCE&G's large customers. Typically they will know if a
11 customer is planning to expand its operations or if a new large customer is requesting
12 service. Finally, when the overall territorial sales forecast is assembled it is reviewed once
13 more to see if it is logical and in line with expectations.

14 **Q. HOW ACCURATE HAS THE SHORT RANGE FORECAST METHODOLOGY**
15 **BEEN AT SCE&G?**

16 **A.** Over the past ten years, 1993-2002, the mean absolute percentage error (MAPE) of the
17 short range forecast was just over 1% comparing actual weather-normalized sales to the
18 projections developed for that year's budget. The largest error of -2.2% occurred in 2001,
19 and was primarily due to the economic slowdown which took place in the second half of the
20 year. The smallest error was -0.2% in 1996.

21 **Q. DOES THIS CONCLUDE YOUR TESTIMONY?**

22 **A.** Yes.

Short-Term Forecasting Groups, 2003 – 2004

<u>Class Number</u>	<u>Class Name</u>	<u>Rate/SIC Designation</u>	<u>Comment</u>
		Single Family	Rates 1, 2, 5, 6, 8, 18, 25, 26, 62, 64
10	Residential Non-Space Heating	Multi Family	Rates 67, 68, 69
910	Residential Space Heating	Mobile Homes	Rates 1, 2, 5, 7, 8
20	Commercial Non-Space Heating	Rate 9	Small General Service
		Rate 12	Churches
		Rate 20, 21	Medium General Service
		Rate 22	Schools
		Rate 24	Large General Service
		Other	Rates 10, 11, 14, 16, 17, 18, 24, 25, 26, 29, 60, 62, 64, 67, 68, 69
920	Commercial Space Heating	Rate 9	Small General Service
30	Industrial Non-Space Heating	Rate 9	Small General Service
		Rate 20, 21	Medium General Service
		Rate 23, SIC 22	Textile Mill Products
		Rate 23, SIC 24	Lumber, Wood Products, Furniture and Fixtures (SIC Codes 24 and 25)
		Rate 23, SIC 26	Paper and Allied Products
		Rate 23, SIC 28	Chemical and Allied Products
		Rate 23, SIC 30	Rubber and Miscellaneous Products
		Rate 23, SIC 32	Stone, Clay, Glass, and Concrete
		Rate 23, SIC 33	Primary Metal Industries; Fabricated Metal Products; Machinery; Electric and Electronic Machinery, Equipment and Supplies; and Transportation Equipment (SIC Codes 33-37)
		Rate 23, SIC 91	Executive, Legislative and General Government (except Finance)
		Rate 23, SIC 99	Other or Unknown SIC Code*
		Rate 27, 60	Large General Service
		Other	Rates 25 and 26
930	Industrial Space Heating	Rate 9	Small General Service
60	Street Lighting	Rates 3, 9, 13, 17, 25, 26, 29, and 69	
70	Other Public Authority	Rate 3 and 29	
		Rates 65 and 66	
92	Municipal	Rate 60, 61	Four Individual Accounts
97	Cooperative	Rate 60, 61	Four Individual Accounts

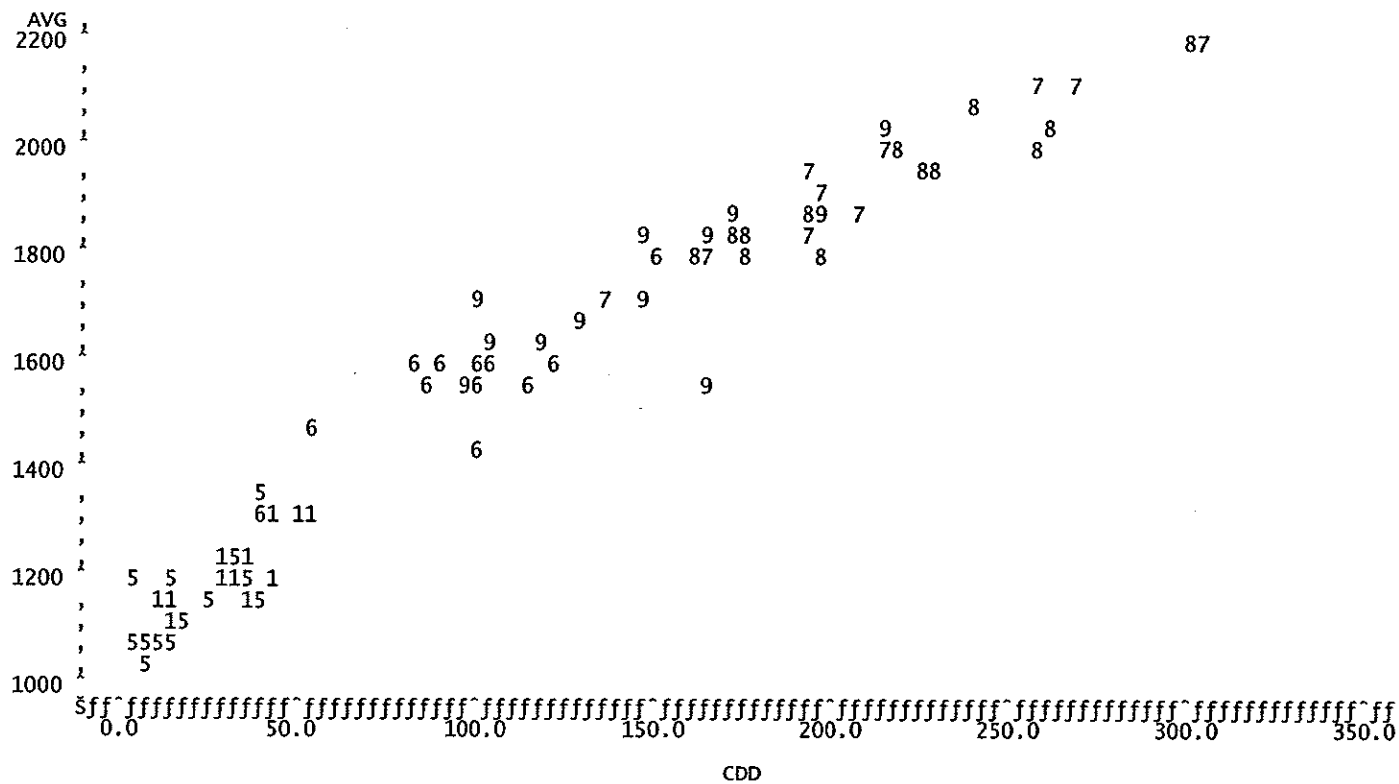
Includes small industrial customers from all SIC classifications that were not previously forecasted individually.

Note: Industrial Rate 23 also includes Rate 24. Commercial Rate 24 also includes Rate 23.

Summary of Methodologies Used To Produce
2003 and 2004 Short Range Forecast

<u>Value Forecasted</u>	<u>Methodology</u>	<u>Forecasting Groups</u>
Average Use	Regression	Class 10, All Groups Class 910, All Groups Class 20, Rates 9, 12, 20, 22, 24, 99 Class 920, Rate 9 Class 70, Rate 3
Total Usage	ARIMA/ Regression	Class 30, Rates 9, 20, 99, and 23, for SIC = 91 and 99 Class 930, Rate 9 Class 60 Class 70, Rates 65, 66
	Regression	Class 92, All Accounts Class 97, All Accounts
Customers	ARIMA	Class 10, All Groups Class 910, All Groups Class 20, All Rates Class 920, Rate 9 Class 30, All Rates Except 60, 99, and 23 for SIC = 22, 24, 26, 28, 30, 32, 33, and 91 Class 930, Rate 9 Class 60 Class 70, Rate 3

MAY - OCTOBER
SINGLE FAMILY HOMES SPACE HEATING
Plot of AVG*CDD. Symbol is value of mo.



MAY - OCTOBER
SINGLE FAMILY HOMES SPACE HEATING

The REG Procedure
Model: MODEL1
Dependent Variable: AVG

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	9	8764686	973854	678.99	<.0001
Error	70	100398	1434.26232		
Corrected Total	79	8865084			

Root MSE	37.87166	R-Square	0.9887
Dependent Mean	1585.54507	Adj R-Sq	0.9872
Coeff Var	2.38856		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > t	Type I SS	Type II SS
Intercept	1	1012.26153	12.80657	79.04	<.0001	201116253	8960829
CDDLO	1	5.40359	0.17370	31.11	<.0001	7228979	1387962
CDDHI	1	3.04862	0.10465	29.13	<.0001	1356334	1217208
Y8909	1	-240.03647	40.26384	-5.96	<.0001	62178	50974
Y8907	1	131.79419	39.06228	3.37	0.0012	7965.51475	16327
Y9406	1	-128.78835	39.40596	-3.27	0.0017	24314	15320
Y9609	1	83.92447	39.88266	2.10	0.0389	10327	6350.93284
Y990506	1	108.90151	27.91055	3.90	0.0002	30573	21835
SEP	1	32.12583	13.57617	2.37	0.0207	9020.76883	8031.24794
TIME	1	0.95232	0.19279	4.94	<.0001	34995	34995

Durbin-Watson D	1.894
Number of Observations	80
1st Order Autocorrelation	0.032